Structure and Method of a Strained Channel Transistor and a Second Semiconductor Component in an Integrated Circuit

[0001] This application claims the benefit of U.S. provisional Application No. 60/497,819 filed on August 26, 2003, and U.S. provisional application Serial No. 60/495,584 filed on August 15, 2003, which applications are hereby incorporated herein by reference.

CROSS-REFERENCE TO RELATED APPLICATIONS

[0002] The following U.S. patents and/or commonly assigned patent applications are hereby incorporated herein by reference:

Patent or Serial No.	Filing Date	Issue Date	Attorney Docket No.
10/379,033	February 28, 2003	________	TSM03-0050
10/667,871	September 22, 2003	________	TSM03-0553
10/641,813	August 15, 2003	________	TSM03-0554
10/628,020	July 25, 2003	________	TSM03-0555
10/627,218	July 25, 2003	________	TSM03-0556
________	________	________	TSM03-0670

TECHNICAL FIELD

[0003] The present invention relates generally to semiconductor devices, and more particularly, the preferred embodiment relates to strained channel complementary field-effect transistors and methods of manufacture.

BACKGROUND

[0004] Size reduction of metal-oxide-semiconductor field-effect transistors (MOSFET), including reduction of the gate length and gate oxide thickness, has enabled the continued improvement in speed performance, density, and cost per unit function of integrated circuits over the past few decades. To enhance transistor performance further, strain may be introduced in the

transistor channel for improving carrier mobilities. Therefore, strain-induced mobility enhancement is another way to improve transistor performance in addition to device scaling. There are several existing approaches of introducing strain in the transistor channel region.

[0005] In one conventional approach, as described in a paper by J. Welser et al., published at the December 1992 International Electron Devices Meeting held in San Francisco, CA, pp. 1000-1002 and incorporated herein by reference, a relaxed silicon germanium (SiGe) buffer layer is provided beneath the channel region. In such a device, a semiconductor device includes a strained silicon layer formed over and abutting a relaxed SiGe layer, which is formed over and abutting a graded SiGe buffer layer.

[0006] The relaxed SiGe layer has a larger lattice constant compared to relaxed Si, and the thin layer of epitaxial Si grown on the relaxed SiGe will have its lattice stretched in the lateral direction, i.e., it will be under biaxial tensile strain. Therefore, a transistor formed on the epitaxial strained silicon layer will have a channel region that is under biaxial tensile strain. In this approach, the relaxed SiGe buffer layer can be thought of as a stressor that introduces strain in the channel region. The stressor, in this case, is placed below the transistor channel region.

[0007] Significant mobility enhancement has been reported for both electrons and holes in bulk transistors using a silicon channel under biaxial tensile strain. In the above-mentioned approach, the epitaxial silicon layer is strained before the formation of the transistor. But there are concerns about the strain relaxation upon subsequent CMOS processing where high temperatures are used. In addition, this approach is very expensive since a SiGe buffer layer with thickness in the order of micrometers has to be grown. Numerous dislocations in the relaxed SiGe buffer layer exist and some of these dislocations propagate to the strained silicon

layer, resulting in a substrate with high defect density. Thus, this approach has limitations that are related to cost and fundamental material properties.

[0008] In another approach, strain in the channel is introduced after the transistor is formed. In this approach, a high stress film is formed over a completed transistor structure formed in a silicon substrate. The high stress film or stressor exerts significant influence on the channel, modifying the silicon lattice spacing in the channel region, and thus introducing strain in the channel region. In this case, the stressor is placed above the completed transistor structure. This scheme is described in detail in a paper by A. Shimizu et al., entitled "Local mechanical stress control (LMC): a new technique for CMOS performance enhancement," published in pp. 433-436 of the Digest of Technical Papers of the 2001 International Electron Device Meeting, which is incorporated herein by reference.

[0009] The strain contributed by the high stress film is believed to be uniaxial in nature with a direction parallel to the source-to-drain direction. However, uniaxial tensile strain degrades hole mobility while uniaxial compressive strain degrades the electron mobility. Ion implantation of germanium can be used to selectively relax the strain so that the hole or electron mobility is not degraded, but this is difficult to implement due to the close proximity of the n and p-channel transistors.

[0010] Accordingly, what is needed in the art is an improved transistor and method thereof that addresses the above-discussed issues.

SUMMARY OF THE INVENTION

[0011] Preferred embodiments of the present invention teach a strained channel transistor and another component formed on the same semiconductor substrate. In a first embodiment, the other component is a resistor. In another embodiment, the other component is a transistor. In other embodiments, the other component can be other devices.

[0012] In one aspect, the invention teaches a method of forming a conventional resistor and a strained channel transistor on the same substrate using the same process flow. A stressor can be defined as that which introduces strain in the transistor channel region. In prior art, schemes of inducing strain in transistors introduce the strain with a stressor, benefiting transistors of the first conduction type while degrading transistors of the second conduction type.

[0013] In accordance with a preferred embodiment of the present invention, a semiconductor chip comprises a semiconductor substrate in which first and second active regions are disposed. A resistor is formed in the first active region; the resistor including a doped region is formed between two terminals. A strained channel transistor is formed in the second active region. The transistor comprises a first and second stressor formed in the substrate oppositely adjacent a strained channel region.

[0014] In accordance with another preferred embodiment of the present invention, a semiconductor chip is formed in a semiconductor region with a first semiconductor material with a natural lattice constant forming a first and second active regions in the semiconductor region. A gate stack is formed over the second active region and a masking layer is formed over the first active region. After forming the masking layer, at least one recess is formed in a portion of the second active region not covered by the gate stack. A second semiconductor material is grown in the recesses, the second semiconductor material having a second natural lattice constant that is

different than the first natural lattice constant. Source and drain regions are formed in the second active region to form a strained channel transistor. The masking layer is removed and a semiconductor component is formed in the first active region.

[0015] In accordance with another preferred embodiment of the present invention, a semiconductor device is formed in a semiconductor substrate with a first semiconductor material. The substrate includes a first active region having a first gate stack and a second active region having a second gate stack. A film is formed over the first and second active regions and spacers are formed on sidewalls of the second gate stack in the second active region. Source and drain recesses are etched on opposing sides of the second gate stack and are spaced from a channel region by the spacers. A second semiconductor material is grown in the source and drain recesses.

[0016] In accordance with another preferred embodiment of the present invention, a semiconductor device is formed by the means of providing a semiconductor layer that includes a first active region and a second active region. A first gate stack is formed over the first active region and a second gate stack is formed over the second active region. A dielectric film is formed over the first and second active regions and a masking layer is formed over a portion of the dielectric film overlying the second active region. Disposable spacers are formed on sidewalls of the first gate stack by anisotropically etching the dielectric film. First and second recesses are formed in the first active region, and are substantially aligned with the disposable spacer. The first and second recesses are filled with a semiconductor material and the source and drain regions in the second active region adjacent the second gate stack are implanted.

BRIEF DESCRIPTION OF THE DRAWINGS

[0017] For a more complete understanding of the present invention, and the advantages thereof, reference is now made to the following descriptions taken in conjunction with the accompanying drawing, in which:

[0018] Figure 1 shows a conventional resistor formed in a portion of substrate;

[0019] Figure 2 shows a strained channel transistor;

[0020] Figure 3 shows the integration of a strained channel transistor and a conventional resistor;

[0021] Figures 4a-4l show a first embodiment process flow;

[0022] Figure 5 compares a conventional PMOS and a compressive stressed PMOS;

[0023] Figure 6 compares a conventional NMOS and a compressive stressed NMOS;

[0024] Figures 7-12 show combined steps of the second and third embodiment;

[0025] Figures 13-14 show additional steps of the second embodiment; and

[0026] Figures 15-19 show additional steps of the third embodiment.

DETAILED DESCRIPTION OF ILLUSTRATIVE EMBODIMENTS

[0027] The making and using of the presently preferred embodiments are discussed in detail below. It should be appreciated, however, that the present invention provides many applicable inventive concepts that can be embodied in a wide variety of specific contexts. The specific embodiments discussed are merely illustrative of specific ways to make and use the invention, and do not limit the scope of the invention.

[0028] Resistors are commonly used in semiconductor integrated circuits. Resistors are used, for example, in analog and in mixed mode analog and digital circuits. Resistors are also used in input and output circuits as input and output resistors. In addition, resistors are sometimes used as part of an input protection circuit to provide protection of the circuit against electrostatic discharge (ESD) events. In this case, the resistor is used to attenuate the ESD voltage and to absorb and dissipate ESD energy. Large voltages in the order of thousands of volts may appear across the two terminals of the resistor used for ESD applications.

[0029] Resistors in integrated circuits may be formed using a poly-crystalline silicon layer, for example. Resistors in integrated circuits may also be formed on a single-crystalline silicon layer, e.g., resistors may be formed in a portion of the single crystal silicon bulk substrate, or in a portion of a single crystal silicon layer in a silicon-on-insulator substrate. As an example, a resistor 100 formed in a portion of a single crystal silicon substrate 102 is shown in Figure 1. The resistor body 104 is doped with a type opposite the substrate 102, and is defined by an isolation structure 106 such as shallow trench isolation, for example. Current 108 flows through the resistor body 104 between two terminals 110 of the resistor 100, as shown in Figure 1. In the resistor body 104, the current 108 experiences a linear current-voltage relationship characteristically defined as resistance. It is known to one skilled in the art that resistors with a

resistor body comprising a single-crystalline semiconductor have the characteristics of high stability and low noise in comparison with conventional poly-crystalline resistor structures.

[0030] In the preferred embodiment, a structure and method of forming a resistor and a strained channel transistor is provided. Methods of forming such resistors with strained channel transistors are provided.

[0031] Figure 2 shows a strained channel transistor 114 where a first semiconductor material in the channel region 116 is stressed by the placement of a second semiconductor material 118 in a part of the source and drain regions 120. The second semiconductor material may also form part of the channel region 116. The lattice constant of the second semiconductor material varies in relation to the lattice constant of the first semiconductor material such that a strain is placed on the first semiconductor material in the channel region. The second semiconductor material will be hereon referred to as a stressor. The transistor 114 comprising a strained channel region 118 is commonly known as a strained channel transistor. When the lattice constant of the stressor (e.g., $Si_{1-x}Ge_x$) is larger than that of the first semiconductor material (e.g., Si), the stressor results in a compressive strain in the source-to-drain direction of the transistor. When the lattice constant of the second semiconductor material (e.g., $Si_{1-y}C_y$) is smaller than that of the first semiconductor material (e.g., Si), the stressor results in a tensile strain in the source-to-drain direction of the transistor. Details of this strained channel transistor are given in co-pending patent application, Y.-C. Yeo *et al.*, "Strained-channel transistor with a lattice-mismatched zone in the source/drain regions," U.S. Patent Application No. 10/379,033, filed March 4, 2003. (TSMC Disclosure Number TSMC2003-0050), and is incorporated herein by reference.

[0032] In the preferred embodiment, the first semiconductor material is silicon (Si) and the second semiconductor material is silicon-germanium (SiGe or $Si_{1-x}Ge_x$), and the strained channel transistor is a p-channel transistor. The mole fraction x of Ge in SiGe may be in the range of about 0.1 to about 0.9. In another embodiment, where the strained channel transistor is an n-channel transistor, the first semiconductor material is silicon, the second semiconductor material is silicon-carbon (SiC or $Si_{1-y}C_y$), and the mole fraction y of C in SiC may be in the range of about 0.01 to about 0.04. While $Si_{1-x}Ge_x$ and $Si_{1-y}C_y$ may be used as the second semiconductor layer, other semiconductor materials may be used. For example, a semiconductor alloy such as $Si_{1-x-y}Ge_xC_y$ may be used as the second semiconductor layer.

[0033] The first embodiment of the present invention will be described with respect to a specific context, namely a method of integrating a conventional resistor such as the resistor with a strained channel transistor. In Figure 3, a conventional resistor 124 is formed in a portion of the substrate 126 in a first active region 138 defined by isolation regions 130, and a strained channel transistor 132 is formed in another portion of the substrate 126.

[0034] The resistor 124 comprises a doped resistor body 128 through which current 134 flows between two resistor terminals 136. The current 136 flowing in the resistor body 128 experiences a resistance, the value of which is a function of many parameters, e.g., the doping type, doping concentration, layout, and size of the resistor body. The doping type of the doped resistor body 128 is opposite the doping type of the semiconductor region 126 immediately underlying the body 128. For example, the resistor 124 may comprise a p+ doped resistor body 128 formed over an n-type doped region 138. The n-type doped region 138 may be an n-type doped well region or an n-type doped substrate 126. It is understood that the doping types may be reversed, e.g., n+ doped resistor body 128 formed on a p-type doped region 138. The doping

distribution or profile in the resistor body is generally non-uniform, and may have an average doping concentration in the range of 10^{16} to 10^{19} cm^{-3}.

[0035] The resistor body 128 shown in Figure 3 can be defined by isolation structures 130, such as shallow trench isolation (STI) structures, for example. The resistor 124 of the present invention may have a rectangular layout with width *W* and a length *L*. The width *W* may have a dimension of larger than about 0.1 microns, and preferably larger than about 1 micron. In the preferred embodiment, the length *L* may have a dimension of larger than about 0.1 micron, and preferably larger than about 1 micron. The resistor may have a layout with a serpentine shape, or any other shape commonly used in the art for diffusion resistors.

[0036] The example in Figure 3 illustrates a bulk semiconductor substrate 126, preferably a bulk silicon substrate. However, it is understood that other substrates such as semiconductor-on-insulator (SOI) substrates may also be used. For example, the semiconductor-on-insulator substrate can be a silicon-on-insulator substrate having a silicon layer overlying a silicon oxide layer, said silicon oxide layer overlying a substrate. The silicon layer in the silicon-on-insulator substrate may be a relaxed silicon layer or a strained silicon layer.

[0037] A cross-section of the resistor 124 in Figure 3 shows a doped body region 128, also known as a resistor body, formed on a portion of the substrate 126. The resistor body 128 can be defined by isolation structures, such as the shallow trench isolation structures 130 shown in Figure 3. The doping type of the doped body region 128 is opposite to the doping type of the semiconductor region 138 immediately underlying the body region 128. For example, if the resistor body 128 is doped p-type, it may be formed on an n-type well region or on an n-type substrate. The average doping concentration of the resistor body 128 may be in the range of 10^{16}

to 10^{19} cm^{-3}. A conductive material can be formed to provide contacts 136 to the terminals of the resistor 124.

[0038] The strained channel transistor 132 of Figure 3 comprises source and drain regions 140 on opposing sides of a channel region 164. The channel region 164, formed from a first semiconductor material 126, is covered by an overlying gate dielectric 150. A gate electrode 148 overlies the gate dielectric 150. The gate electrode 148 material can be poly-crystalline silicon, poly-crystalline silicon germanium, metals, metallic silicides, metallic nitrides, or conductive metallic oxide. Spacers 170 consisting of one or more dielectric materials are formed on the sidewalls of the gate electrode 148. A portion of the source and drain regions 140 comprise a second semiconductor material 162. The second semiconductor material 162 may have a second natural lattice constant that is different from the natural lattice constant of the first material 126. A silicide 174 overlays the gate electrode 148 and the source and drain regions 140. In contrast, the doped region constituting the resistor body 128 is not silicided to maintain a high resistance.

[0039] Principles of the present invention can also be applied to a resistor of the type taught in co-pending application Serial No. 10/667,871, filed September 22, 2003 (TSM03-0553), which application is incorporated herein by reference. Using the methods taught herein, this resistor can be formed simultaneously with a strained channel transistor.

[0040] The present invention teaches a method of forming the strained channel transistor 132 on the same semiconductor substrate 126 as the conventional resistor 124 using the same fabrication or manufacturing process.

[0041] Referring now to Figure 4a, a process flow showing the method of manufacturing a resistor with a strained channel transistor is described. A semiconductor substrate 126, preferably a silicon substrate, is provided and isolation structures 130 are formed to define active

regions in the substrate. The isolation structures 130 may be formed using standard shallow trench isolation processes, for example, comprising the steps of etching trenches with depths in the range of about 2000 to about 6000 angstroms, and filling the trenches with a trench filling dielectric material by chemical vapor deposition, for example, to give the cross-section as shown in Figure 4a. The trench filling dielectric may be silicon oxide, for example. Ion implantation may be performed to form n-type and/or p-type well regions (not shown). Figure 4a shows two active regions: a first active region 142 where a conventional resistor 124 is to be formed, and a second active region 144 where a strained channel transistor 132 is to be formed. These active regions might be of the same conductive type as each other or they may be of different conductivity types. Source/drain regions 140 are shown in the Figure 4a even though these regions have not been formed yet.

[0042] A gate stack 146 is then formed in the second active region 144, as shown in Figure 4b. The gate stack 146 comprises a gate electrode 148 overlying a gate dielectric 150. The gate stack 146 may additionally comprise a gate mask 152 overlying the gate electrode 148. The purpose of incorporating the gate mask 152 will become clear below.

[0043] The gate stack may be formed by the following process. A gate dielectric 150 is formed in the second active region 144 using any gate dielectric formation process known and used in the art, e.g., thermal oxidation, nitridation, sputter deposition, or chemical vapor deposition. The physical thickness of the dielectric 150 may be in the range of about 5 to about 100 angstroms. The transistor gate dielectric 150 may employ a gate dielectric such as silicon oxide and silicon oxynitride or a high permittivity (high-k) gate dielectric, or combinations thereof.

[0044] The high-k dielectric preferably has a permittivity of larger than 8. This dielectric can be one or more of aluminum oxide (Al_2O_3), hafnium oxide (HfO_2), hafnium oxynitride (HfON), hafnium silicate ($HfSiO_4$), zirconium oxide (ZrO_2), zirconium oxynitride (ZrON), zirconium silicate ($ZrSiO_4$), yttrium oxide (Y_2O_3), lanthalum oxide (La_2O_3), cerium oxide CeO_2), titanium oxide (TiO_2), tantalum oxide (Ta_2O_5), or combinations thereof. In the preferred embodiment, the high-k dielectric is hafnium oxide. The silicon equivalent oxide thickness (EOT) of the dielectric 150 is preferably less than about 50 angstroms, more preferably less than about 20 angstroms, and even more preferably less than about 10 angstroms. The physical thickness of the dielectric 150 may be less than about 100 angstroms, more preferably less than about 50 angstroms, and even more preferably less than about 20 angstroms.

[0045] After the gate dielectric 150 is formed, a gate electrode material 148 can then be deposited over the gate dielectric 150 layer. The gate electrode material 148 can be poly-crystalline silicon, poly-crystalline silicon germanium, metals, metallic silicides, metallic nitrides, or conductive metallic oxide. In the preferred embodiment, the electrode 148 comprises poly-crystalline silicon. Metals such as molybdenum, tungsten, titanium, tantalum, platinum, and hafnium may be used as the portion of the top electrode 148. Metallic nitrides may include, but are not restricted to, molybdenum nitride, tungsten nitride, titanium nitride, and tantalum nitride. Metallic silicides may include, but will not be restricted to, nickel silicide, cobalt silicide, tungsten silicide, titanium silicide, tantalum silicide, platinum silicide, and erbium silicide. Conductive metallic oxides may include, but will not be restricted to, ruthenium oxide and indium tin oxide.

[0046] The gate electrode material 148 may be deposited by conventional techniques such as chemical vapor deposition. The gate electrode 148 may also be formed by the deposition of

silicon and metal, followed by an annealing to form a metal silicide gate electrode material. A patterned gate mask 152 is then formed on the gate electrode 148 material using conventional deposition and photolithography techniques. The gate mask 152 may employ commonly used masking materials such as, but not limited to, silicon oxide, silicon oxynitride, and silicon nitride. The gate electrode 148 is then etched using plasma etch processes to form the gate electrode. The gate dielectric 150 on regions not covered by the gate electrode 148 is preferably etched away.

[0047] As shown in Figure 4c, a first mask material 154 is deposited over the gate stack 146. The first mask material 154 may be a dielectric such as silicon oxide, silicon oxynitride, or silicon nitride, for example. In the preferred embodiment, the first mask material comprises a silicon nitride on silicon oxide multi-layer.

[0048] A second mask material 156 is then formed using deposition and photolithographic techniques to cover the first mask material 154 in the first active region 142, while exposing the first mask material 154 in the second active region 144 as shown in Figure 4d. The second mask material 156 may comprise any masking material that is different from the first mask material 154. In the preferred embodiment, the second mask material 156 comprises a photoresist.

[0049] An etching of the first mask material 154 in the second active region 144 is then performed in the presence of the second mask material 156. The etching is preferably an anisotropic etch done using plasma etching techniques. This results in spacers or liners 158 being formed adjacent to the gate stack 146 in the second active region 144, as shown in Figure 4e. The second mask material 156 may be removed at this point.

[0050] A recess with depth *d* is etched in the source and drain regions, as shown in Figure 4f. The etch may be accomplished by a plasma etch using chlorine and bromine chemistry. The

depth *d* of the recess may range from about 50 angstroms to about 1000 angstroms. An optional anneal may be performed to facilitate silicon migration to repair any etch damage as well as to slightly smoothen the silicon surface for the subsequent epitaxy process.

[0051] Next, a second semiconductor material 162 is epitaxially grown to at least partially fill the recessed region 160. This can be accomplished using selective epitaxial growth. The epitaxy process used to perform the epitaxial growth may be chemical vapor deposition, ultra-high vacuum chemical vapor deposition (UHV-CVD), or molecular beam epitaxy. The epitaxially grown materials may also extend above the surface of the channel region 164 of the transistor 132, forming a raised source and drain structure (not shown). In the first preferred embodiment, the second semiconductor material 162 comprises of silicon germanium with a germanium mole fraction between about 0.1 and about 0.9. In the second preferred embodiment, the lattice-mismatched zone is comprised of silicon-carbon with a carbon mole fraction of between about 0.01 and about 0.04.

[0052] The gate mask 152 covers the top portion of the gate electrode 148 so that no epitaxial growth occurs on the gate electrode 148. The liner 158 covers the sidewalls of the gate electrode 148 so that no epitaxial growth occurs on the sidewalls. Epitaxial growth on the gate electrode 148 sidewalls potentially results in an electrical short between the gate stack and the source and drain regions 140.

[0053] An optional cap layer may be epitaxially grown to cover the second semiconductor material 162. For example, the optional cap layer may comprise a first semiconductor material 126, as shown in Figure 4g. The purpose of having the cap layer is to facilitate the subsequent formation of a low resistance silicide in the source and drain regions 140.

[0054] Following epitaxial growth, the gate mask 152 can be removed. The liner 158 can be optionally removed.

[0055] The epitaxially grown first and second semiconductor materials, 126 and 162 respectively, may be in-situ doped or undoped during the epitaxial growth. If undoped as grown, they may be doped subsequently and the dopants activated using a rapid thermal annealing process. The said dopants may be introduced by conventional ion implantation, plasma immersion ion implantation (PIII), gas or solid source diffusion, or any other techniques known and used in the art. Any implant damage or amorphization can be annealed through subsequent exposure to elevated temperatures. A first shallow implant can be first performed to dope the shallow regions of the resistor body 128 and to form the source/drain extensions, 140 of the transistor 132, as shown in Figure 4h.

[0056] A spacer 170 is then formed, followed by a second and deeper implant. The second implant additionally dopes the resistor body 128, and also forms the deep source and drain regions 140 of the strained channel transistor 132. The structure formed at this stage is shown in Figure 4i.

[0057] The resistance of the source and drain in the transistor can be reduced by strapping the source/drain regions 140 with a silicide 174, e.g., using a self-aligned silicide (salicide) process, or other metal deposition process. This is illustrated in Figure 4j. A mask, usually comprising an oxide, is typically used prior to the silicidation process to cover portions of the substrate where silicidation is not intended. For example, the oxide mask covers the first active region 142 while exposing the second active region 144. A subsequent silicidation process therefore forms silicides 174 on the gate electrode 148, and source and drain regions 140 of the strained channel transistor 132, while no silicide is formed in the first active region 142 where

the resistor 124 is located. While not shown, resistor 124 contacts can be formed by the silicidation process.

[0058] Next, a contact etch stop layer 176 may be formed, followed by the deposition of a passivation layer 178, as shown in Figure 4k. Contact holes 180 are then etched through the passivation layer 178, stopping on the contact etch stop layer 176. A conductive material is then filled into the contact holes 180 to form conductive contacts to the resistor 124 and the strained channel transistor 132 as shown in Figure 4l.

[0059] In the first embodiment, a resistor and strained channel transistor are integrated into a single device. In the next embodiment, a strained channel transistor is integrated into the same chip as a non-strained channel transistor. Since the use of a contact etch stop over the non-strained transistor could result in strain, in this context, a non-strained channel transistor is meant to include a transistor that is not strained using source/drain stressors.

[0060] The second embodiment will be described in the context of an integration flow is described for manufacturing an improved CMOS device. As before, the source and drain regions are etched and then refilled of silicon, geranium, carbon, or combinations thereof. The alloy is deposited on the layer of silicon by a selective epitaxy process thereby creating a stress in the channel of the transistor between the source and drain. The larger lattice spacing creates a compressive stress and the smaller one creases a tensile stress.

[0061] Compressive stress improves carrier mobility of the PMOS transistor and degrades carrier mobility of the NMOS transistor, as shown in Figures 5 and 6. It is an objective of certain embodiments of this invention to separate n-channel and p-channel transistors by engineering the nature and magnitude of the strain in the channel region of the transistors. It is desirable to induce a compressive strain in the channel of the p-channel transistor in the source-

to-drain direction and compressive stress free of the n-channel transistor. It is also desirable to induce a tensile strain in the channel of the n-channel transistor in the source-to-drain direction and tensile stress free of the p-channel transistor.

[0062] Another preferred embodiment of the present invention teaches a method of integrating strained channel transistors of more than one conduction type with minimal degradation of carrier mobility.

[0063] Referring now to Figure 7, a process flow showing the method of manufacturing strained channel transistors of multiple conduction types with minimal degradation of carrier mobility is described. A semiconductor substrate 200, preferably a silicon substrate, is provided and isolation structures 202 are formed to define active regions in the substrate. The isolation structures 202 may be formed using standard shallow trench isolation (STI) processes, for example, comprising the steps of etching trenches with depths in the range of about 2000 to about 6000 angstroms, and filling the trenches with a trench filling dielectric material by chemical vapor deposition to give the cross-section as shown in Figure 7. The trench filling dielectric 202 may be silicon oxide, for example. Ion implantation may be performed to form n-type well regions 204 or p-type well regions 206. Figure 7 shows two active regions: a first active region 208 where a p-type strained channel transistor is to be formed, and a second active region 210 where an n-type channel transistor is to be formed.

[0064] A gate stack 212 is then formed in the first and second active regions 208/210, as shown in Figure 7. The gate stack 212 may additionally comprise a hard mask 218 overlying the gate electrode 214. The gate stack 212 comprises a gate electrode 214 overlying a gate dielectric 216. The gate dielectric 216 is formed using any gate dielectric formation process known and used in the art, e.g., thermal oxidation, nitridation, sputter deposition, or chemical vapor

deposition. The physical thickness of the gate dielectric 216 may be in the range of 5 to 100 angstroms. The gate dielectric 216 may employ a conventional gate dielectric such as silicon oxide and silicon oxynitride or a high permittivity (high-k) gate dielectric, or combinations thereof.

[0065] The high-k dielectric preferably has a permittivity of larger than 8. This dielectric can be one or more of aluminum oxide (Al_2O_3), hafnium oxide (HfO_2), hafnium oxynitride (HfON), hafnium silicate ($HfSiO_4$), zirconium oxide (ZrO_2), zirconium oxynitride (ZrON), zirconium silicate ($ZrSiO_4$), yttrium oxide (Y_2O_3), lanthalum oxide (La_2O_3), cerium oxide CeO_2), titanium oxide (TiO_2), tantalum oxide (Ta_2O_5), or combinations thereof. In the preferred embodiment, the high-k dielectric is hafnium oxide. The silicon equivalent oxide thickness (EOT) of the dielectric 150 is preferably less than about 50 angstroms, more preferably less than about 20 angstroms, and even more preferably less than about 10 angstroms. The physical thickness of the dielectric 150 may be less than about 100 angstroms, more preferably less than about 50 angstroms, and even more preferably less than about 20 angstroms.

[0066] After the gate dielectric 216 is formed, a gate electrode material 214 can then be deposited over the gate dielectric 216. The gate electrode material 214 can be comprised of poly-crystalline silicon, poly-crystalline silicon germanium, metals, metallic silicides, metallic nitrides, or conductive metallic oxide. In the present embodiment, the electrode 212 comprises poly-crystalline silicon. Metals such as molybdenum, tungsten, titanium, tantalum, platinum, and hafnium may be used as the portion of the top electrode 214. Metallic nitrides may include, but are not restricted to, molybdenum nitride, tungsten nitride, titanium nitride, and tantalum nitride. Metallic silicides may include, but will not be restricted to, nickel silicide, cobalt silicide, tungsten silicide, titanium silicide, tantalum silicide, platinum silicide, and erbium

silicide. Conductive metallic oxides may include, but will not be restricted to, ruthenium oxide and indium tin oxide.

[0067] The gate electrode material 214 may be deposited by conventional techniques such as chemical vapor deposition. The gate electrode 214 may also be formed by the deposition of silicon and metal, followed by an annealing to form a metal silicide gate electrode material. A patterned hard mask 218 is then formed on the gate electrode 214 material using deposition and photolithography techniques. The gate mask 218 may employ commonly used masking materials such as, but not limited to, silicon oxide, silicon oxynitride, and silicon nitride. The gate electrode 214 is then etched using plasma etch processes to form the gate electrode. The gate dielectric 216 on regions not covered by the gate electrode 214 is preferably etched away.

[0068] As shown in Figure 8, a disposable film 220 is formed over the first and second active regions 208/210. The disposable film may be a dielectric film formed using a chemical vapor deposition process or sputter deposition. The disposable film may comprise oxide, for example. In the preferred embodiment, the disposable film 220 is between about 10 and about 1000 angstroms thick, and more preferably between about 10 and about 200 angstroms thick.

[0069] A first mask material 222 shown in Figure 9 is deposited over the first and second active regions 208/210 may be silicon oxide, silicon oxynitride, or silicon nitride. In the preferred embodiment, the first mask material comprises a silicon nitride on a silicon oxide multi-layer.

[0070] Figure 10 shows a second mask material 224 formed over the second active region 210 using deposition and photolithographic techniques to cover the first mask material 222 in the second active region 210, while exposing the first mask material 222 in the first active region 208 of Figure 10. The second mask material 224 may comprise any masking material that is

different from the first mask material 222. In the preferred embodiment, the second mask material comprises a photoresist.

[0071] An etching of the first mask material 222 in the second active region 210 is then performed in the presence of the second mask material 224. The etching is preferably an anisotropic etch done using plasma etching techniques. This results in disposable spacers or liners 226 being formed adjacent to the gate stack 212 in the first active region 208, as shown in Figure 11.

[0072] After the disposable spacers 226 are formed, recessed regions 228 are etched in the active area substantially aligned with the disposable spacers 226. A silicon etch chemistry can be used as discussed above. The second mask material 224 may be removed after etching.

[0073] Next, as shown in Figure 12, second semiconductor material 230 is epitaxially grown to at least partially fill the recessed region 228. This can be accomplished using selective epitaxial growth (SEG). The epitaxy process used to perform the epitaxial growth may be chemical vapor deposition (CVD), ultra-high vacuum chemical vapor deposition (UHV-CVD), or molecular beam epitaxy (MBE). The epitaxially grown materials may also extend above the surface of the channel region 232 of the second active region 210, forming a raised source and drain 230 structure as shown in Figure 12. In the second preferred embodiment, the second semiconductor material 230 comprises of silicon germanium with a germanium mole fraction between about 0.1 and about 0.9. In the second preferred embodiment, the lattice-mismatched zone is comprised of silicon-carbon with a carbon mole fraction of between about 0.01 and about 0.04.

[0074] The gate mask 218 covers the top portion of the gate electrode 214 so that no epitaxial growth occurs on the gate electrode 214. The disposable liner 226 prevents epitaxial growth on the gate electrode 214 sidewalls.

[0075] Following epitaxial growth, the gate mask 218, disposable liner 226, and first mask material can be removed, forming the structure shown in Figure 13.

[0076] The epitaxially grown first 200 semiconductor materials may be in-situ doped or undoped during the epitaxial growth. If undoped as grown, it may be doped subsequently and the dopants activated using a rapid thermal annealing process. The dopants may be introduced by conventional ion implantation, plasma immersion ion implantation (PIII), gas or solid source diffusion, or any other techniques known and used in the art. Any implant damage or amorphization can be annealed through subsequent exposure to elevated temperatures.

[0077] Figure 14 shows the semiconductor device after further processing. A first shallow implantation can be performed on the structure of Figure 14 to dope the shallow regions of the first and second transistor source and drain regions and to form the source/drain extensions, as shown in Figure 14.

[0078] Spacers (including regions 244 and 246) are formed on the sides of the gate electrode 214. In one example, the spacers may be formed by chemical vapor deposition of a dielectric material, e.g., silicon oxide or silicon nitride, followed by an anisotropic etching of the dielectric material to form simple spacers. In the example of Figure 14, the spacers are composite spacers. A composite spacer may comprise a dielectric liner 244 and a spacer body 246. The dielectric liner 244 may be formed by the deposition of a dielectric liner material, e.g., silicon oxide, and the spacer body material 246, e.g. silicon nitride, followed by an anisotropic etch using reactive

ion etching. In another embodiment, the liner 244 may be an oxide and the spacer body 246 may be a nitride.

[0079] The source and drain regions for the first transistor 236 are formed using ion implantation while covering the second transistor 234. In the preferred embodiment, the dopant is arsenic or phosphorus or a combination of both. The source and drain regions for the second transistor 234 formed by using ion implantation while covering the first transistor 236. In the preferred embodiment, a dopant such as boron is used. A passivation layer 248 is formed over the first and second active regions 208/210.

[0080] A third embodiment of the present invention will now be described with respect to Figures 15-19. Figure 15 shows the structure of Figure 12 after further processing. In particular, a source/drain implantation step has been performed as described above. In this case, the implanted dopants extend through the second semiconductor material 230 into the first well region 204. In this case, the source/drain regions include second semiconductor material 230 as well as the doped portion 240 of the first semiconductor material 200.

[0081] A third protective layer 252 shown in Figure 16, preferably a photoresist, is then formed using deposition and photolithographic techniques to cover the first active area 208 while exposing the second active area 210. An etching of the first mask material 222 in the second active region 210, as described above, results in disposable spacers 226 being formed adjacent to the gate stack 212 in the second active region 210, as shown in Figure 16.

[0082] Doped regions 240 in the first semiconductor material 200 are formed using doping methods described above. Any implant damage or amorphization can be annealed through subsequent exposure to elevated temperatures. Following a deep implant and the removal of the spacers 226 of the first and second transistors 236/234, an additional shallow implant can be

performed to dope the source and drain extension regions 238of the first and second transistors 236/234. The resulting structure is shown in Figure 17.

[0083] Figure 18 shows the semiconductor device after further processing. Additional steps may include forming a liner 244 and a spacer 246 on the sides of the gate stacks 212 for the first and second transistors 236/234, and forming an etch stop layer 248 covering the first and second transistors 236/234.

[0084] Figure 19 shows an alternate embodiment where the spacers 244/246 have been eliminated. One purpose of the spacers in an embodiment such as shown in Figure 14, for example, is to mask the source/drain extensions (e.g., lightly doped drain) during formation of the heavily doped source and drain region. As shown in Figures 16 and 17, however, the heavily doped source and drain regions 240 are formed prior to the formation of extensions 238. Accordingly, the spacers are not needed for this purpose. In another embodiment not shown, spacers or other sidewall lines can be included that do not align with the heavily doped source and drain regions 240.

[0085] The resistance of the gate, source and drain of the first and second transistors 236/234 can be reduced by strapping the gate electrode 214, and source and drain regions 230/240 with a silicide 250, e.g., using a self-aligned silicide (salicide) process, or other metal deposition process. These silicided regions are shown in Figure 18.

[0086] In the two embodiments described, a strained channel transistor is formed in the same substrate as a resistor and another transistor. In another embodiment, all three components can be formed in the same substrate.

[0087] In other embodiments, other components can be formed with the strained channel transistor. For example, a capacitor is described in a pending application Serial No. 10/627,218, filed July 25, 2003 (TSM03-0556). In another example, a diode or lubistor is described in a co-pending application Serial No. 10/628,020, filed July 25, 2003 (TSM03-0555). Both of these applications are incorporated herein by reference. Using the concepts taught herein, any of the structures taught in the co-pending applications can be formed in the same substrate as the strained channel transistor.

[0088] In the foregoing specification, the invention has been described with reference to specific embodiments. However, various modifications and changes can be made by one skilled in the art without departing from the scope of the preferred embodiment. Accordingly, the specification and figures are to be regarded in an illustrative rather than a restrictive sense, and all such modifications are intended to be included within the scope of the preferred embodiment.

[0089] Moreover, the scope of the present application is not intended to be limited to the particular embodiments of the process, machine, manufacture, composition of matter, means, methods and steps described in the specification. As one of ordinary skill in the art will readily appreciate from the disclosure of the present invention, processes, machines, manufacture, compositions of matter, means, methods, or steps, presently existing or later to be developed, that perform substantially the same function or achieve substantially the same result as the corresponding embodiments described herein may be utilized according to the present invention. Accordingly, the appended claims are intended to include within their scope such processes, machines, manufacture, compositions of matter, means, methods, or steps.